Using impartial combinatorial games as a benchmark for adaptable learning algorithms

Carnegie Mellon University

Alp Müyesser[1,2]
1. Department of Mathematical Sciences
2. Center for the Neural Basis of Cognition
Carnegie Mellon University, Pittsburgh, PA

Kyle Dunovan[3,4]
3. Department of Psychology
4. Center for the Neural Basis of Cognition
University of Pittsburgh, Pittsburgh, PA

Timothy Verstynen[5,6]
5. Department of Psychology
6. Center for the Neural Basis of Cognition
Carnegie Mellon University, Pittsburgh, PA

CNBC CENTER FOR THE NEURAL BASIS OF COGNITION

Background & Motivation

- Artificial learning agents aspire to match and exceed the human brain in challenging contexts
- Since benchmarks selected favor *accuracy* (e.g. beating the current grandmaster) over *transfer-learning* (e.g. how do I get good at a lot of chess-variants simultaneously), trained networks are inflexible to changes in inputs and goals.

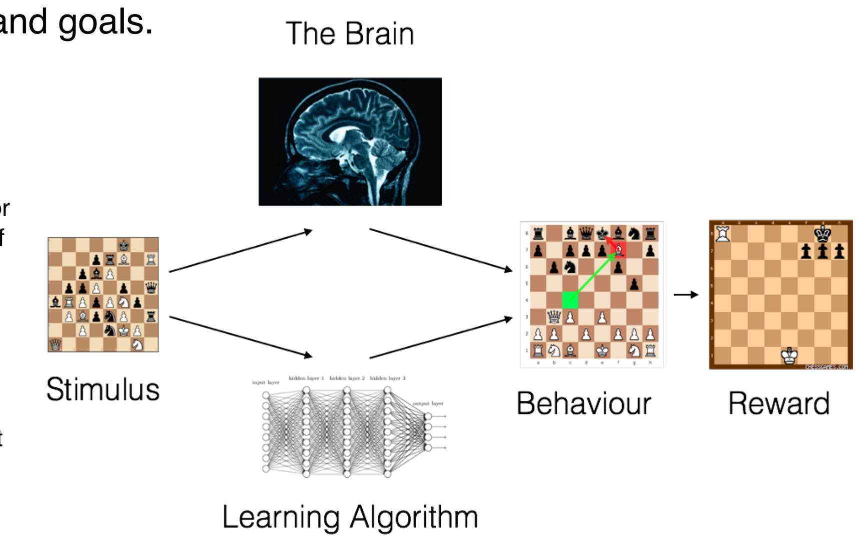


Figure 1. Learning tasks under the brain as the information-processor framework. A learning algorithm attempts to fill in for the brain given the context of a game. The context of the task chosen shapes how the learning algorithm is designed; DeepMind and AlphaGo are very distinct in their architecture since they tackle different problems, yet neither concerns itself with efficiency across multiple tasks.

Impartial Combinatorial Games

Impartial combinatorial games offer numerous advantages when taken as a benchmark for a potential learning agent that strives to achieve *transfer-learning*:

- Impartial games are simple to analyze, and often have mathematical structure that is "easy to discover".
- The rules of impartial games immediately generalize to bigger board sizes.
- Numerous ways to manipulate the rules of the game while preserving general strategies

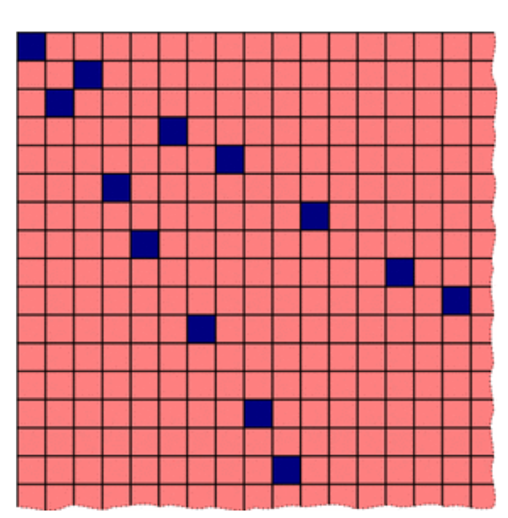


Figure 2. *Wythoff's Game* visualized. Coordinates in the grid where the origin is the upper-left corner are the game positions. A valid move for P1 or P2 is a horizontal, vertical, or a diagonal move towards the origin. The winner is the player who makes the move to the origin.

If the game is in a red position, player 1 can win under optimal play.

Abstractions in Biological/Artificial Networks

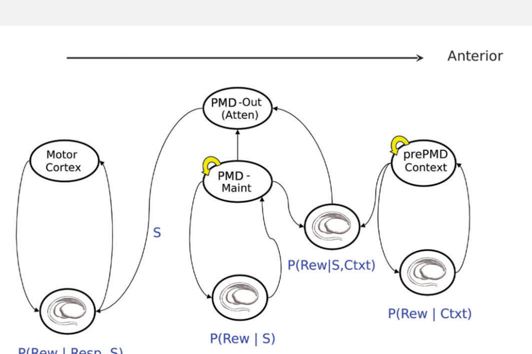


Figure 3. Schematic of hierarchical corticostraital unit. (Frank & Badre)

- It is proposed that corticostraital loops are organized along a gradient of abstractions, and this framework allows humans to relate similar tasks to transfer experience. (Frank & Badre)
- Networks have been designed with similar structures in visual recognition learning tasks (Huang et. al.)

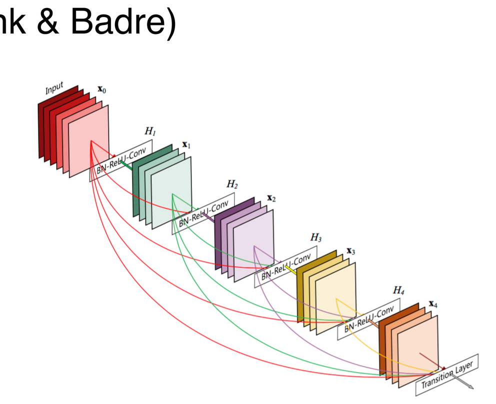
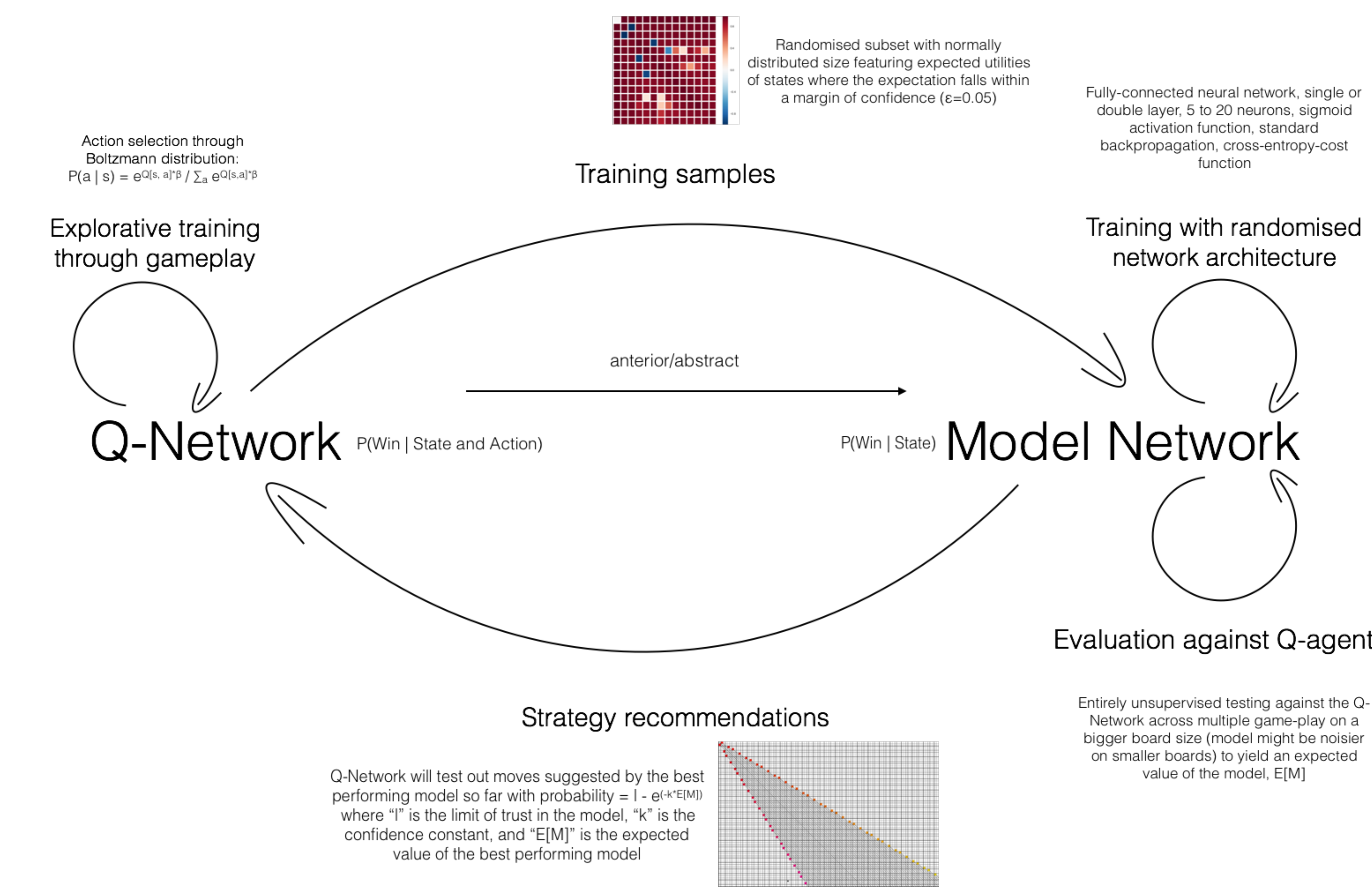


Figure 4. Convolutional neural network with hierarchical layers (Huang et. al.)

Deep Model Network (DML)

A *hybrid network* architecture that prioritizes model-building through navigating the possible model space using exploratory Q-network as a performance measure

- The *Model Network* is trained using a noisy random sample from the Q-Network featuring expected values of the game positions from the Q-Network
- The Q-Network continues exploration given strategy recommendations from the Model Network until an optimal model that vastly outperforms the Q-Network on varying board sized is obtained



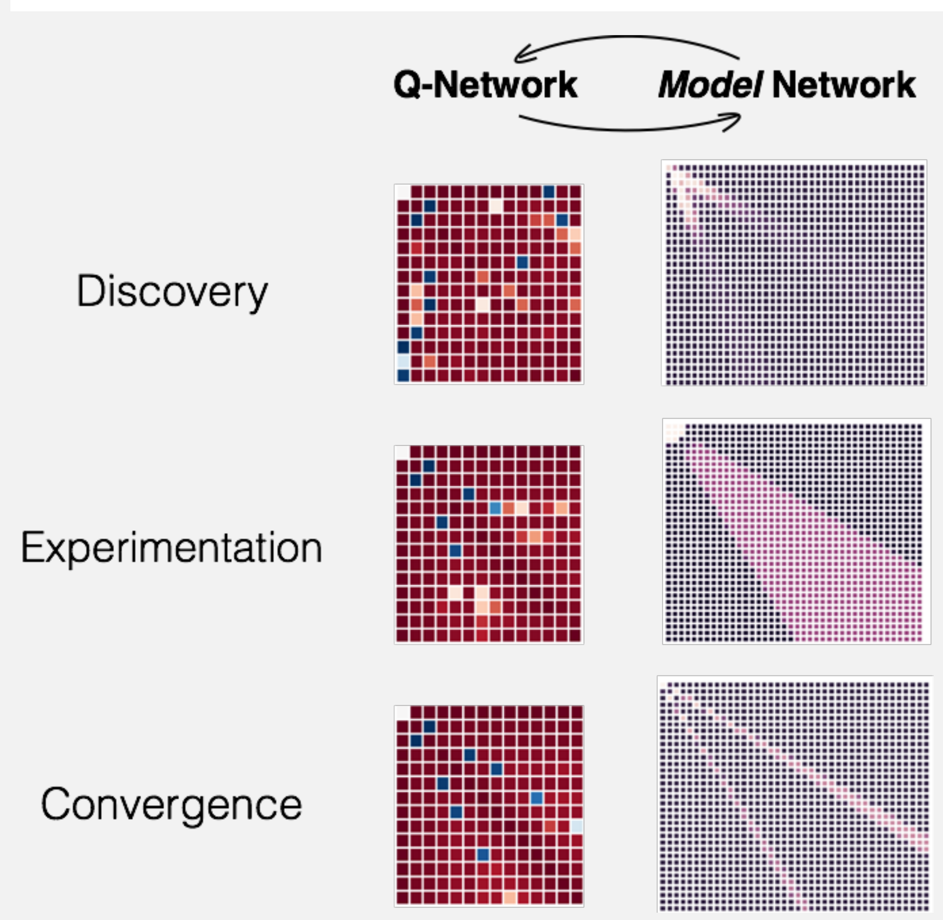


Figure 5. The DML in various stages of training. The early models (Discovery) will be largely unsuccessful, while certain inaccurate generalizations (Experimentation) will supply reasonable strategies to the Q-Network, allowing the provision of useful datasets into the model network that translate into accurate and general models (Convergence)

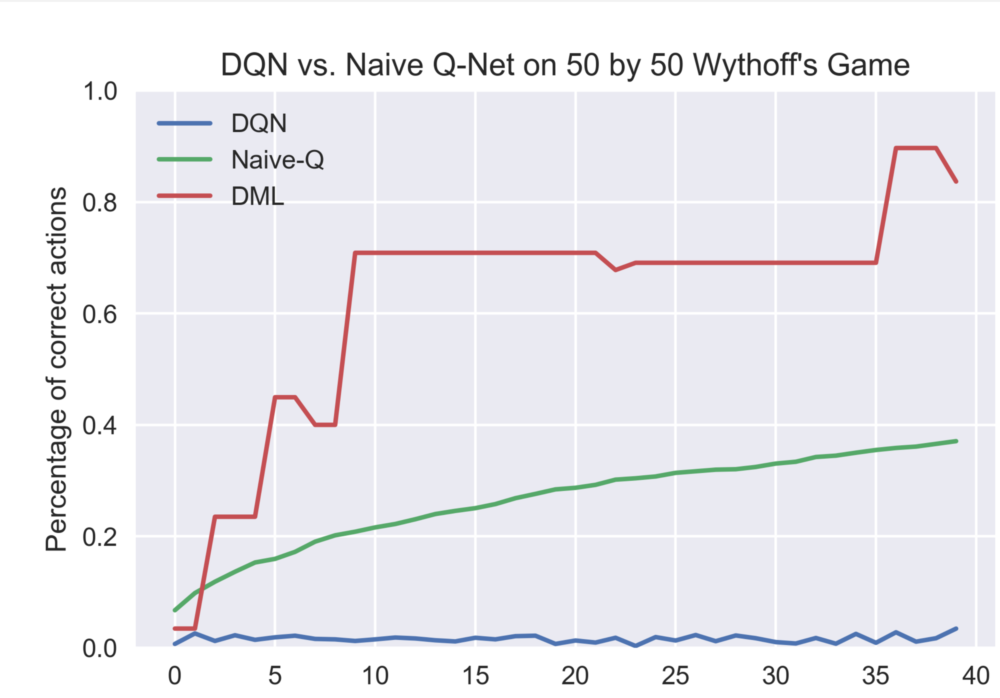


Figure 6. The DML vastly outperforms the Naïve-Q and the DQN on identical training periods. The Naïve-Q is trained on a 50 by 50 board on 5000 gameplay simulations per time-step, whereas the DQN is trained on a 12 by 12 board with 1000 simulations each. The DML learns better models in discrete jumps whereas the Naïve-Q has a steady but diminishing learning rate. DQN (single-layer, standard backpropagation, fully-connected, sigmoid activation function, sum-squared-error cost function, lr=0.01) and Naïve Q-Net (y=1, lr=0.1, β =0.6) compared. Time-steps calculated on a 2.7 GHZ Intel i5 2-core processor (2015 MacBook Pro).

Performance across board sizes

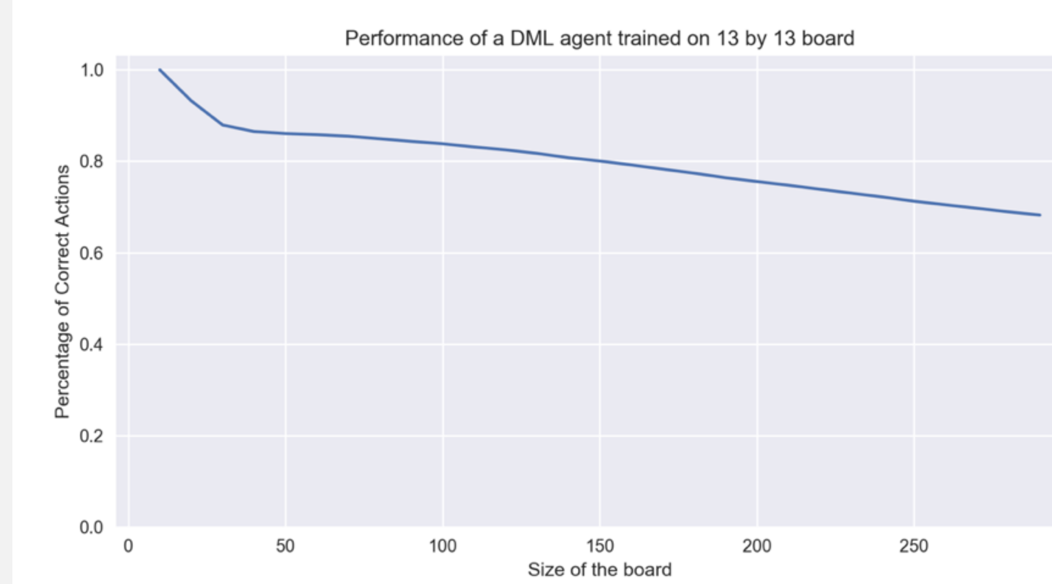


Figure 6. Performance of a Model Network that the DML agent converged to through datasets received from a Q-Network trained on a 13 by 13 board. Single layer feedforward network with cross-entropy cost function, trained for 2000 epochs.

- The DML not only outperforms the existing learning agents in terms of efficiency, but also is able to generalize experience on to bigger board sizes.

Performance across Wythoff-Variants

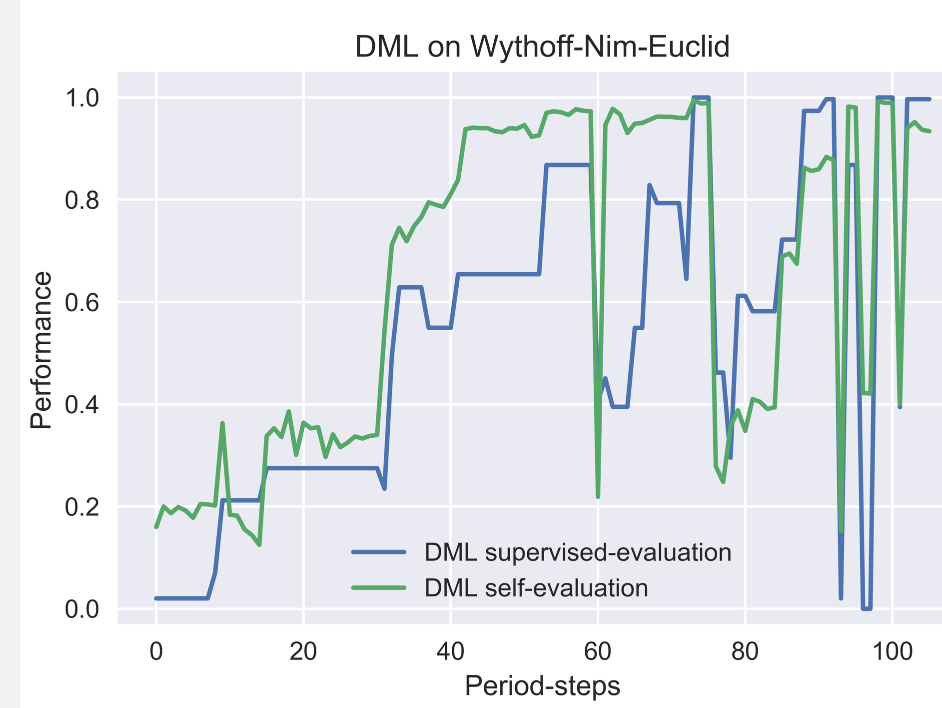


Figure 6. Performance of a DML across 3 different impartial combinatorial games. When the DML converges on a model with performance exceeding a certain threshold, the rules of the game are changed. The DML discovers this, and adapts and reuses old models if they apply to the new set of rules. Performance decreases become less drastic as DML learns all three games simultaneously.

- The DML is able to converge to near-optimal performance even when the rules of the game are being changed, by training different models for different games, and reusing old models when they are applicable.

Conclusion / Future Directions

- The flexible nature of the context of impartial combinatorial games allowed us to devise DML, which greatly outperforms standard machine learning approaches in the field.
- The DML could generalize experience to bigger board sizes, and it could adapt to various modifications to the rules of the game, however, it does not transfer-learn across games that have similar shape and structure, indicating that an extension of DML featuring a Symbolic layer to relate the abstractions of games with similar rules can be useful

Acknowledgements

Frank, M. & Badre, D. (2012) Mechanisms of Hierarchical Reinforcement Learning in Corticostraital Circuits 1: Computational Analysis. *Cerebral Cortex* 509-26.

Huang, G., Liu, Z., Weinberger, K. & Maaten, L. (2016) Densely Connected Convolutional Neural Networks. arXiv: 1608.06993

*This work was funded by a Center of the Neural Basis of Cognition grant (PNC Fellowship)